



UNIVERSIDAD DE GRANADA

TRABAJO FIN DE GRADO

INGENIERÍA INFORMÁTICA

Aprendizaje Automático para la extracción de características y detección de situaciones anómalas en multitudes

Autor

Diego Navarro Cabrera

Directores

Name of the main supervisor

Name of the second supervisor (if available)



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA Y
TELECOMUNICACIONES

—
Granada, 22 de abril de 2021

**Aprendizaje Automático para la extracción de características y
detección de situaciones anómalas en multitudes**

Diego Navarro Cabrera

Palabras clave:

Resumen

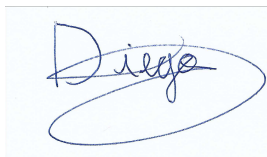
**Machine learning for feature extraction and abnormal crowd
behavior**

Diego Navarro Cabrera

Keywords:

Abstract

Yo, **Diego Navarro Cabrera**, alumno del Grado de Ingeniería Informática de la **Escuela Técnica Superior de Ingeniería Informática y Telecomunicaciones de la Universidad de Granada**, con DNI 75935043Z, autorizo la ubicación de la siguiente copia de mi Trabajo Fin de Grado en la biblioteca del centro para que pueda ser consultada por las personas que lo deseen.

A handwritten signature in blue ink, reading "Diego", enclosed within a large, loopy oval stroke.

Fdo: Diego Navarro Cabrera

Granada, 22 de abril de 2021

D. **Name of the main supervisor** y D.^a **Name of the second supervisor** (if available), profesores del Departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada.

Informan:

Que el presente trabajo, titulado *Aprendizaje Automático para la extracción de características y detección de situaciones anómalas en multitudes*, ha sido realizado bajo su supervisión por **Diego Navarro Cabrera**, y autorizamos la defensa de dicho trabajo ante el tribunal que corresponda.

Y para que conste, expiden y firman el presente informe en Granada a 22 de abril de 2021

Los directores:



Name of the main supervisor



Name of the second supervisor
(if available)

Agradecimientos

Índice general

Índice general	5
1 Introducción	6
2 Fundamentación Teórica	7
2.1. Aprendizaje automático	7
2.2. Máquina de soporte vectorial	7
2.3. Extracción de características	7
2.4. Autoencoders	7
2.5. Aprendizaje multi-tarea	7
3 Descripción del problema	8
3.1. Conjuntos de datos de trabajo	8
3.1.1. UMN Crowd Dataset	9
3.1.2. Violent Flows	10
3.2. Estado del arte	10
3.2.1. Modelo basado en descriptores visuales locales de nivel medio	10
3.3. Modelo propuesto	10
4 Resultados experimentales	11
4.1. Metodología usada	11
4.1.1. UMN Crowd Dataset	11
4.1.2. Violent Flows	11
4.2. Comparación de resultados	11
4.2.1. UMN Crowd Dataset	11
4.2.2. Violent Flows	11
5 Conclusión	12
Bibliografía	13

1

Introducción

Intro

Fundamentación Teórica



-
- 2.1. Aprendizaje automático
 - 2.2. Máquina de soporte vectorial
 - 2.3. Extracción de características
 - 2.4. Autoencoders
 - 2.5. Aprendizaje multi-tarea

Descripción del problema

3

En este capítulo desarrollaremos en más profundidad el problema a tratar. Como ya hemos mencionado anteriormente, nuestro objetivo es estudiar el comportamiento de una multitud para detectar cuándo se está produciendo una anomalía.

En concreto nos centraremos en el enfoque holístico o *top-down* de este problema. Este enfoque trata a la multitud como un solo ente y analiza la dinámica del grupo más que las acciones de sus individuos. La detección de comportamientos individuales anómalos dentro de la multitud queda fuera del ámbito de este trabajo pero algunas propuestas de las estudiadas podrían adaptarse para ello.

La estructura de esta sección será la siguiente. Primero hablaremos de los conjuntos de datos sobre los que trabajaremos. A continuación trataremos el estado del arte y explicaremos en más profundidad el modelo sobre el que nos hemos basado. Finalmente detallaremos el modelo desarrollado y diversas alternativas que se han tenido en cuenta

3.1. Conjuntos de datos de trabajo

Se han elegido 2 conjuntos de datos ampliamente usados en el ámbito de la detección de anomalías en multitudes. El primero es el UMN Crowd Dataset [1], que consiste de una pequeña selección de vídeos en el que un grupo de personas hecha a correr tras una señal. El segundo dataset es conocido como Violent Flows y recoge una mezcla de vídeos de peleas y vídeos de control. Estos dos conjuntos de datos nos permitirán comprobar la efectividad del modelo ante 2 de las anomalías más habituales y relevantes en el día a día: las peleas y la fuga de personas. Además su amplio uso en la literatura nos permitirá comparar fielmente nuestro modelo con otros muchos . A continuación describiremos más detalladamente ambos conjuntos.

3.1.1. UMN Crowd Dataset

Este conjunto de datos, elaborado por la Universidad de Minesota (UMN), consiste de 11 vídeos en los que un grupo de personas se comporta de manera normal hasta que recibe una señal y todos se dispersan corriendo. Los 11 vídeos se dividen en 3 escenas distintas, cada una con 2, 6 y 3 vídeos respectivamente. Todos estos vídeos se encuentran unidos en un mismo archivo que conviene dividir antes de ser procesado.



Figura 3.1: Ejemplo de frame normal (izqda) y anómalo (drcha) en la escena 1.

Este dataset clasifica los frames normales y los anómalos por medio de un cartel que aparece en pantalla cuando se produce la anomalía, tal y como se puede ver en 3.1. Para evitar que esto pueda influir en nuestro modelo se ha recortado la parte superior de cada vídeo para eliminar dicho cartel. Por otro lado, tal y como se detalla en [2] estos carteles no son exactos, ya que aparecen significativamente después del comienzo de la anomalía. Es por esto que se ha decidido usar las anotaciones provistas en dicho artículo para etiquetar el comienzo de esta en cada vídeo. Además de esto, en la mayoría de los casos la anomalía acaba antes de que empiece el siguiente vídeo, por lo que también tendremos que anotar el frame de finalización de cada caso. Hablaremos con más detalle de esto en el apartado de experimentación.

A pesar de sus ventajas y su uso muy extendido, este dataset presenta algunos inconvenientes que hay que tener en cuenta. Para empezar, su tamaño de 11 vídeos es bastante limitado, más aún si tenemos en cuenta que están divididos en 3 escenas que han de evaluarse de manera separada. Además solo representa anomalías de un tipo muy específico (gente echando a correr), por lo que no nos sirve para evaluar modelos enfocados a detectar todo tipo de anomalías, algo de lo que hablaremos cuando propongamos nuestro modelo.

3.1.2. Violent Flows

3.2. Estado del arte

3.2.1. Modelo basado en descriptores visuales locales de nivel medio

3.3. Modelo propuesto

Resultados experimentales

4

4.1. Metodología usada

4.1.1. UMN Crowd Dataset

4.1.2. Violent Flows

4.2. Comparación de resultados

4.2.1. UMN Crowd Dataset

4.2.2. Violent Flows

5

Conclusión

Conclusión

Bibliografía

-
- [1] *UMN Crowd Dataset*. http://mha.cs.umn.edu/proj_events.shtml#crowd/.
 - [2] Duan-Yu Chen y Po-Chung Huang. “Motion-based unusual event detection in human crowds”. En: *J. Visual Communication and Image Representation* 22 (feb. de 2011), págs. 178-186. doi: [10.1016/j.jvcir.2010.12.004](https://doi.org/10.1016/j.jvcir.2010.12.004).